DOCUMENT RESUME

ED 268 011 SE 046 539

AUTHOR Callman, Joshua L.; And Others
TITLE Computer-Based Semantic Network in Molecular Biology: A Demonstration.
PUB DATE Oct 85
NOTE 6p.; Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (59th, San Francisco, CA, March 28-April 1, 1986). For related papers presented in this same set, see SE 046 537-540.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Biology; Cognitive Psychology; *College Science; *Computer Oriented Programs; *Computer Software; *Concept Formation; *Concept Teaching; Higher Education; Knowledge Level; Science Education; Science Instruction
IDENTIFIERS *Concept Mapping; Science Education Research

ABSTRACT

This paper analyzes the hardware and software features that would be desirable in a computer-based semantic network system for representing biology knowledge. It then describes in detail a prototype network of molecular biology knowledge that has been developed using Filevision software and a Macintosh computer. The prototype contains about 100 linked concept maps. Design principles are delineated. For example, the decision was made not to include direct links between members of a set such as "RNA" and "DNA." To move from one to the other in the network it is necessary to move up the hierarchy to "nucleic acids" and back down. This requirement is pedagogically attractive and reduces clutter. Many other such principles are required to assure consistency in map generation. Working with information in the computer-based semantic network provides an experience heretofore unavailable. Project members find it is fun, and somewhat challenging even for creators of the database, to find paths from point X to point Y in the database. Moving through the net requires an integration of concepts that seems qualitatively different from that encountered in most other interactions with academic subject matter. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED268011

III. Computer-Based Semantic Network in Molecular Biology: A Demonstration

Joshua L. Callman, Joseph Faletti, and Kathleen M. Fisher

Graduate Group in Science and Mathematics Education
University of California — Berkeley
Berkeley, CA 94720

and

Departments of Education and Computer Science
University of California — Davis
Davis, CA 95616

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

October 1985

Abstract

This paper analyzes the hardware and software features that would be desirable in a computer-based semantic network system for representing biology knowledge. It then describes in detail a prototype network of molecular biology knowledge that has been developed using Filevision software and a Macintosh computer. The prototype contains about one-hundred linked concept maps.

Design principles are delineated. For example, the decision was made not to include direct links between members of a set such as "RNA" and "DNA." To move from one to the other in the network it is necessary to move up the hierarchy to "nucleic acids" and back down. This requirement is pedagogically attractive and reduces clutter. Many other such principles are required to assure consistency in map generation.

Working with information in the computer-based semantic network provides an experience heretofore unavailable. Project members find it is fun, and somewhat challenging even for creators of the database, to find paths from point X to point Y in the database. Moving through the net requires an integration of concepts that seems qualitatively different from that encountered in most other interactions with academic subject matter.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Kathleen M. Fisher

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SE 046 539

BEST COPY AVAILABLE

III. Computer-Based Semantic Network in Molecular Biology: A Demonstration

Joshua L. Callman, Joseph Faletti, and Kathleen M. Fisher

Graduate Group in Science and Mathematics Education
University of California — Berkeley
Berkeley, CA 94720

and

Departments of Education and Computer Science
University of California — Davis
Davis, CA 95616

October 1985

1. Purpose

The purposes of this paper are fourfold: To summarize the desired features of a computer-based semantic network of biological knowledge; to describe (and demonstrate) a prototype semantic network of molecular biology knowledge that has been developed on a Macintosh computer using Filevision software; to test (with students and faculty) those features of the network which can be implemented in the prototype; and to identify, to the extent possible, the emergent features of a semantic network (i.e., those properties which characterize the net but not the smaller units of information that comprise the net).

2. Theoretical Basis

Background theory has been developed in previous papers. Here we will consider the ideal features of a computer-based semantic network of biology knowledge.

Choice of Hardware/Software. In constructing pedagogical software, it is important that the resulting software will run on machines which are available in the schools. Two pilot prototypes of the semantic network were constructed, aimed at the lowest common denominator of currently available machines, one on an Apple II-E and one on the DEC PDP/VAX families of minicomputer. These were used to assess how students interact with information which is presented in various non-graphic fashions and whether such representations might be used in instruction. Our findings forced reconsideration of the minimum hardware/software capabilities required for this project. Since the capabilities of commonly available and affordable machines are rapidly expanding, our revised requirements (described below) do not seem unreasonable.

The computer should have good graphics capabilities, allowing easy construction of concept maps. A relatively large high-resolution screen is highly desirable (certainly larger than that on a Mac). A mouse or similar pointing device should be available to avoid large amounts of typing by users. The computer should be able to

BEST COPY AVAILABLE

read graphic information (as with Thunderscan software on the Mac) and to print out graphic data. The computer should be very fast with lots of memory.

Constructor Editor. A program is needed which facilitates drawing concept maps on the computer and linking them together into a network. The program should be so easy to use that undergraduate students can easily begin generating their own maps, once they develop sufficient knowledge of biology and of mapping strategies. It should have such features as the following. The constructor should contain heuristics governing graphic generation that lead to appropriate placement of the various parts of a concept map and their interconnections. It should allow "zooming in" and "zooming out", with automatic adjustments in the amount and kind of information presented at close and far views of a network. It should be easy to enter new information into the net and to change existing information, either concepts or relationships. When a new concept is entered, the computer should prompt the user with an appropriate set of relationships. Menus should be available summarizing all available relationships and features. Directions should be available in a well-indexed help file.

Browser. It should be easy to browse, wandering through the network simply by clicking the desired relation or concept with a mouse or mouse-like pointer. It would be nice to be able to retrieve a summary of the path taken.

Research Tool. The computer-based semantic network could be a powerful research tool, providing insights into the thought processes of different individuals. It should be possible to gather information about what choices a user makes and what paths they take. Such information can then be used, for example, to compare the activities of experts with novices, or to observe changes in strategies used by a particular student at two different points in time.

Images. Many biological concepts are intimately tied to images. It is not clear how imagery and verbal knowledge are related in the mind, but it is clear that there is a close interconnection. Thus, we are considering incorporating images as well as words into the semantic network of biological knowledge.

Games. Interesting pedagogical tasks (games) can be provided to motivate the student to use the network. A variety of game structures have been designed in preliminary form. One type of task we have used is to instruct students to find a path through the network from concept A to concept B, as in "Trace energy from the sun to a coyote." The game becomes more challenging if certain information, such as the relations, are masked. In order to move from one concept to the next, the student must indicate the nature of the relationship between the two concepts. Such a game structure could theoretically "keep score," counting the number of moves and number of errors. It would provide a summary of the path taken at the end of the game, and might compare this with paths taken by others (the last player, the instructor, the longest & shortest paths, etc.). The game structures generally require additional manipulative capabilities of the software for accessing the network.

BEST COPY AVAILABLE

3. Procedure

A system that provides some of the desired features is the Macintosh computer with Filevision software. Filevision is a commercially available Macintosh program that provides a set of tools for creating graphics (similar to but different from MacPaint), allows differential highlighting of certain features of the graphics, and allows linking together of frames. A prototype Filevision-based semantic network of molecular biology knowledge was produced. It consists of about a hundred interlinked Filevision "documents." Each document contains a central concept in molecular biology along with its constellation of closely related concepts (e.g., Figure 1).

With this system, we were able to fit about 60 maps on a single disk. By creating an internal RAM disk for the Filevision program, the two external disks could be used for concept maps, nearly doubling the capacity of the system (we do not have a hard disk). Each concept map was linked via Filevision facilities to all other relevant concept maps. In this prototype, it is possible to move through the network along the myriad of available paths, simply by pointing with the mouse and clicking the button. Concepts or relations can be masked and unmasked, allowing some approximation of the desired game environment, although the process is clumsy.

A number of design principles have been established to standardize the arrangement of relations and related concepts around the central concept. Related concepts which are conceptually or structurally part of the central concept are placed below. Related concepts which are larger than the central concept, either structurally or conceptually, are place above it. In concept maps representing processes, inputs are on the left and outputs are on the right.

4. Results

Positive Outcomes. The Filevision prototype is the first we have created in which it is relatively easy to enter data and organize a knowledge base. It is also the first to give the user a feeling of moving through a complex net. In some ways it is so convincing that we have difficulty making users realize how limited the system really is; they keep forgetting (or failing to understand) that this is simply a prototype, and imagine that the system is capable of much more than it actually is. The notion of assembling linked concept maps appears to be a sound one. Just finding paths from concept "X" to concept "Y" is fun, and somewhat challenging, even for those who helped create the net.

Negative Outcomes. The Filevision software is slow and clumsy. It makes us aware of how important it will be to have a fast machine. The Filevision system is highly constrained, lacking most of the features of the ideal network beyond the basic notion of linked concept maps. Drawing maps and making changes are tedious processes. There are no heuristics or similar tools for speeding up map construction. Yet while constraining in some ways, the Filevision software is not constraining enough in other ways -- for example, it does not require or enforce consistency or completeness. Great self-discipline is necessary to completely enter all the appropriate connections and to use names consistently.

Emergent Properties. The network does have some nice emergent properties. Moving through the net requires an integration of concepts that seems qualitatively

BEST COPY AVAILABLE

different from that encountered in most other interactions with academic subject matter. We found that the systematic organization of knowledge in network form is as useful for the instructor (in instructional design and task analysis) as it is for the student (in studying). These uses will be illustrated in detail in the final paper.

Design Decisions. Theoretically, all concepts in biology can be related to all other concepts. When is a direct link between two concepts to be included in the net, and when not? Some clear design decisions have been made. For example, members of a set (such as "RNA" and "DNA") are not connected directly to one another by a member-member relationship. To move from "RNA" to "DNA", it is necessary to move up the hierarchy to "nucleic acid" and back down again. Two reasons dictated this decision: the first, pragmatic, to minimize clutter on the screen; the second, pedagogical, to promote the user's facility with hierarchical organization of knowledge. A full set of such design guidelines will be presented. Many decisions remain to be made, particular regarding connections between distant points in the net.

5. Implications for Teaching

It is possible to teach about the cell in many different ways and present the same information in many different sequences. Irrespective of instructional approach, representation of this knowledge in semantic network form shows precisely where bridges can be built between topics. This helps teachers to convey new information to students in ways that allow the students to connect the information to what they already know. Semantic networks provide a precise scaffolding which can be used to tease apart the myriad of complex interconnections which exist in biological knowledge.

BEST COPY AVAILABLE